Advanced Quantitative Methods

Instrumental Variables

Instructor: Gregory Eady
Office: 18.2.10
Office hours: Fridays 13-15

Today

- Instrumental variables
- Exercise

Fields experiments are the gold standard for causal inference, but:

- Practical challenges
  - Money
  - Time
  - Access
  - Event already happened
- Ethical problems

If we can't run a field experiment, how can we estimate a causal effect in a real-world setting?

- **Instrumental variables (IV)**
  - An external actor or nature causes as-if random variation in a treatment of interest
- **Regression discontinuity design (RDD)**
  - Treatment is assigned by a rule-based threshold (e.g. you can vote on your 18th birthday onward, but not the day before)
- **Differences-in-differences (DD)**
  - Treatment status varies among units across time (e.g. some states legalize marijuana in the last 10 years, others do not)

Some classic instrumental variables examples

The Colonial Origins of Comparative Development: An Empirical Investigation

By Daron Acemoglu, Simon Johnson, and James A. Robinson*

We exploit differences in European mortality rates to estimate the effect of institutions on economic performance. Europeans adopted very different colonization policies in different colonies, with different associated institutions. In places where Europeans faced high mortality rates, they could not settle and were more likely to set up extractive institutions. These institutions persisted to the present. Exploiting differences in European mortality rates as an instrument for current institutions, we estimate large effects of institutions on income per capita. Once the effect of institutions is controlled for, countries in Africa or those closer to the equator do not have lower incomes. (*JEL* O11, P16, P51)

Why are some countries rich and others poor?

- Can we find a *causal* explanation for the big drivers of income differences?
- Acemoglu et al. (2001) hypothesize that "strong" institutions *cause* long-term economic development
- Weak (colonial) institutions are extractive:
    - Little protection of private property
    - No checks and balances on government
    - Small colonial footprint
- Strong (colonial) institutions mimic Europe:
    - Protection of private property
    - Checks and balances in government
    - Larger settler colonies (US, Canada, Australia)

Problem

- The strength of institutions will be due to a large number of factors that are themselves causes of economic development
- We cannot simply try to find all of the controls
  - No one will really believe that you've captured all of the unobserved confounders or somehow dealt with reverse causality
- So how can we get exogenous variation in the strength of institutions?
- Seems impossible. How can institutions be randomly assigned?

Solution: Mosquitoes (settler mortality)

- Some colonies had higher mortality rates due to yellow fever and malaria
- Thus Europeans:
  - Created *weak* extractive institutions where Europeans were more likely to die from disease
  - Created *strong* property-protecting institutions where European were less likely to die from disease
- The consequence: The disease environment *as-if* randomly assigned weak institutions to some places, and strong institutions to others

Empowering Women Through Radio: Evidence from Occupied Japan*

Yoko Okuyama †

August, 2023

Abstract

I study the impact of women's radio programs that the US-led occupying force aired in Japan (1945-1952) to dismantle the prewar patriarchal norms. Through the lens of economics of identity, the radio messages can be viewed as attempts to alter gendered identity norms, and thus to shift women's outcomes. Using local variation in radio signal strength driven by soil conditions as an instrumental variable, I show that greater exposure to women's radio programs increased women's electoral turnout, and the vote share for female candidates, highlighting women's votes matter. Moreover, exposure to women's radio programs accelerated the postwar fertility decline.

Does news media tailored toward women increase women's representation in politics?

- Does exposure to radio programs for women increase turnout and vote share for women candidates? (context is post-war Japan)
- Problem: The type of women who listen to women's radio will be different from those who do not?
- How do we get exogenous variation in exposure to women's radio programs?

Solution: Soil quality

- Exposure depends on ground wave field strength, which depends on the type of soil and its salt/moisture
- Areas will certain soil types will have patchier coverage than areas with other soil types
- The consequence: Soil type *as-if* randomly assigns radio quality to potential subscribers (and thus the number of subscribers)

DO POLITICAL PROTESTS MATTER? EVIDENCE FROM THE TEA PARTY MOVEMENT*

ANDREAS MADESTAM
DANIEL SHOAG
STAN VEUGER
DAVID YANAGIZAWA-DROTT

Can protests cause political change, or are they merely symptoms of underlying shifts in policy preferences? We address this question by studying the Tea Party movement in the United States, which rose to prominence through coordinated rallies across the country on Tax Day, April 15, 2009. We exploit variation in rainfall on the day of these rallies as an exogenous source of variation in attendance. We show that good weather at this initial, coordinating event had significant consequences for the subsequent local strength of the movement, increased public support for Tea Party positions, and led to more Republican votes in the 2010 midterm elections. Policy making was also affected, as incumbents responded to large protests in their district by voting more conservatively in Congress. Our estimates suggest significant multiplier effects: an additional protester increased the number of Republican votes by a factor well above 1. Together our results show that protests can build political movements that ultimately affect policy making and that they do so by influencing political views rather than solely through the revelation of existing political preferences. *JEL* Code: D72.

Do protests cause an increase political support in an election?

- Did the 2009 Tea Party protests increase support for Republicans?
- Simple question, but how can we get a causal answer?

Problem

- Protest turnout in a given city will be a function of existing support for Republican candidates
- Using regression with controls probably won't be believable
- How do we get exogenous (i.e. as-if random) variation in protests?
- Seems impossible because protests are not randomly located

Solution: Rainfall

- Whether it rains on the day of a protest can be thought of as-if random
- If it rains, a protest might get cancelled in some places (or at least fewer people will attend)
- The consequence: Rainfall as-if randomly assigns protests to some places, but not to others

When it rains, it pours...

Poverty and Witch Killing

EDWARD MIGUEL
University of California, Berkeley and NBER

First version received March 2003; *final version accepted February* 2005 (*Eds*.)

This study uses rainfall variation to estimate the impact of income shocks on murder in rural Tanzania. Extreme rainfall (drought or flood) leads to a large increase in the murder of "witches"—typically elderly women killed by relatives—but not other murders. The findings provide novel evidence on the role of income shocks in causing violent crime, and religious violence in particular.

Economic Shocks and Civil Conflict: An Instrumental Variables Approach

Edward Miguel

University of California, Berkeley and National Bureau of Economic Research

Shanker Satyanath and Ernest Sergenti

New York University

Estimating the impact of economic conditions on the likelihood of civil conflict is difficult because of endogeneity and omitted variable bias. We use rainfall variation as an instrumental variable for economic growth in 41 African countries during 1981–99. Growth is strongly negatively related to civil conflict: a negative growth shock of five percentage points increases the likelihood of conflict by one-half the following year. We attempt to rule out other channels through which rainfall may affect conflict. Surprisingly, the impact of growth shocks on conflict is *not* significantly different in richer, more democratic, or more ethnically diverse countries.

Shaping the Nation: The Effect of Fourth of July on Political Preferences and Behavior in the United States*

Andreas Madestam† and David Yanagizawa-Drott‡

November 2011

Abstract

This paper examines whether social interactions and cultural practices affect political views and behavior in society. We investigate the issue by documenting a major social and cultural event at different stages in life: the Fourth of July celebrations in the United States during the 20th century. Using absence of rainfall as a proxy for participation in the event, we find that days without rain on Fourth of July in childhood shift adult views and voting in favor of the Republicans and increase turnout in presidential elections. The effects we estimate are highly persistent throughout life and originate in early age. Rain-free Fourth of Julys experienced as an adult also make it more likely that people identify as Republicans, but the effect depreciates substantially after a few years. Taken together, the evidence suggests that political views and behavior derive from social and cultural experience in early childhood, and that Fourth of July shapes the political landscape in the Unites States.

Does the length of incarceration for defendants decrease employment opportunities & future income?

Incarceration Length, Employment, and Earnings

Jeffrey R. Kling

AMERICAN ECONOMIC REVIEW
VOL. 96, NO. 3, JUNE 2006
(pp. 863-876)

Download Full Text PDF

Article Information

Abstract

This paper estimates effects of increases in incarceration length on employment and earnings prospects of individuals after their release from prison. I utilize a variety of research designs including controlling for observable factors and using instrumental variables for incarceration length based on randomly assigned judges with different sentencing propensities. The results show no consistent evidence of adverse labor market consequences of longer incarceration length using any of the analytical methods in either the state system in Florida or the federal system in California. (JEL: J24; K42)

Problem

- Criminals who are incarcerated for long periods are likely much different from those who are incarcerated for short periods
- Using regression with controls probably won't be believable
- How do we get exogenous (i.e. as-if random) variation in incarceration?
- Seems impossible because the law should dictate incarceration length

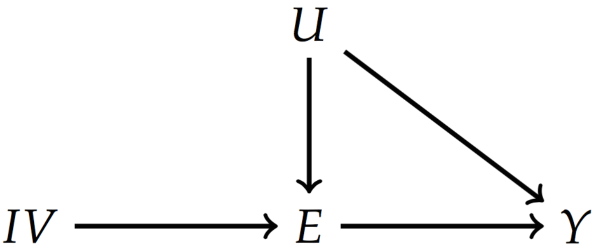
Solution: Judge severity

- Whether a defendant gets a short or long prison sentence depends on how "strict" the judge is who decides the sentence
- In many jurisdictions, defendants are randomly assigned to one of, say, ten judges working during the day that they are to be sentenced
- The consequence: As-if random assignment to a lenient or strict judge determines a defendant's sentence length
- Note: This strategy works for any context in which consequences for a person are decided by a randomly assigned assessor/committee
  - E.g. refugee/immigration boards, insurance claim assessors, patent examiners, graders, doctors deciding treatments

Instrumental variables: The basic idea

- An "exogenous" variable (the instrument) causes as-if random variation in your "endogenous" variable (your variable of interest)
- Importantly, the instrumental variable must *only* affect the outcome *through* your variable of interest
- This is called the *exclusion restriction*, which will discuss a bit more in just one second

IV as a diagram



We then get the following set of connections

- "Reduced form": the effect of your instrumental variable on the outcome
- "First stage": the effect of your instrumental variable on the endogenous variable
- "Second stage": the instrumented effect of your endogenous variable on the outcome

Two-stage least squares (2SLS)

First stage predicts the variation in the endogenous variable X_i that is caused by the instrumental variable Z_i:

$$X_i = \beta_0 + \beta_1 Z_i + \epsilon_i \tag{1}$$

Second stage then calculates the effect of $\hat{X}_i$ (the variation caused by the instrument) on the outcome:

$$Y_i = \delta_0 + \delta_1 \hat{X}_i + \gamma_i \tag{2}$$

Note: the "hat" on $\hat{X}_i$ just denotes the predicted value of X_i as a function of the instrument Z_i (as calculated from Equation (1), the first stage regression)

IV estimates are local average treatment effects (LATE)

- It is *not* the average treatment effect (ATE)
- It is the causal effect on Y_i by the part of X_i that is causally affect by your instrumental variable
- e.g. The effect of being assigned to treatment on the type of people who would open their door for canvassers.
- e.g. The effect of attending a 4th of July celebration on the type of person who would stay or not stay home if it rained
- Important to think about this conceptually when you consider how generalizable your findings are

Assumptions

1. **Independence ("exogeneity")**
   - i.e. there is a causal effect of Z_i on X_i
   - Put differently, Z_i causes as-if random variation in X_i
2. **Relevance ("strong" instrument)**
   - i.e. the instrument is strongly predictive of the endogenous variable (this is *not* about statistical significance)
   - An F-statistic measures how much additional variation in the endogenous variable is explained by the instrument
   - It compares a first-stage model *with* the instrument to a "restricted" model *without* the instrument:
     - $y_i = \alpha + \beta_1 Z_i + \beta_2 X_i + \epsilon_i$ (full model)
     - $y_i = \alpha + \beta_2 X_i + \epsilon_i$ (restricted model)
   - F-statistic > 10 is a rule of thumb (Staiger & Stock 1997)
3. **Monotonicity**...
4. **Exclusion restriction**....

3. Monotonicity

Under two-sided non-compliance:

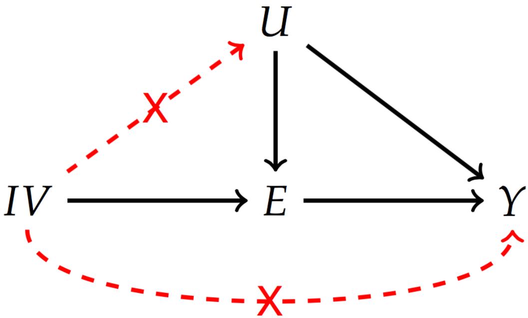
- Compliers
- Never-takers
- Always-takers
- Defiers

Monotonicity means there are no defiers:

- Binary IV: No one who is treated does the opposite of the expected
- Continuous IV: X_i only increases or only decreases as Z_i increases

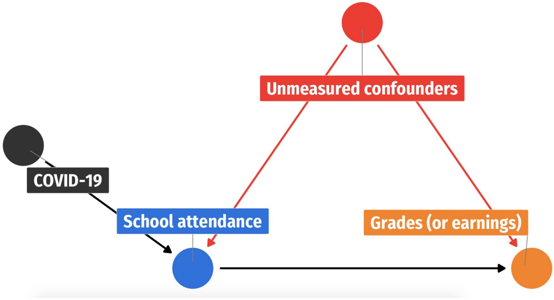
4. Exclusion restriction

Z_i has an effect on Y_i *only* through its effect on D_i (i.e. "E" in the diagram below, which is the endogenous variable)



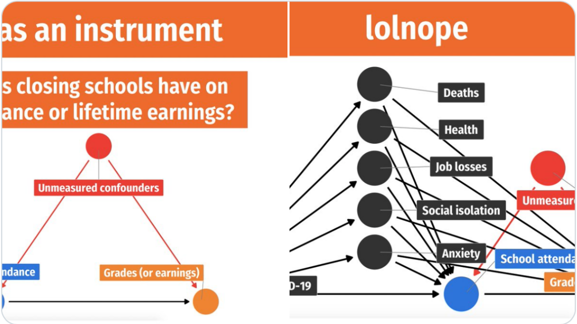
COVID-19 as an instrument

What effect does closing schools have on student performance or lifetime earnings?



Andrew Heiss
@andrewheiss

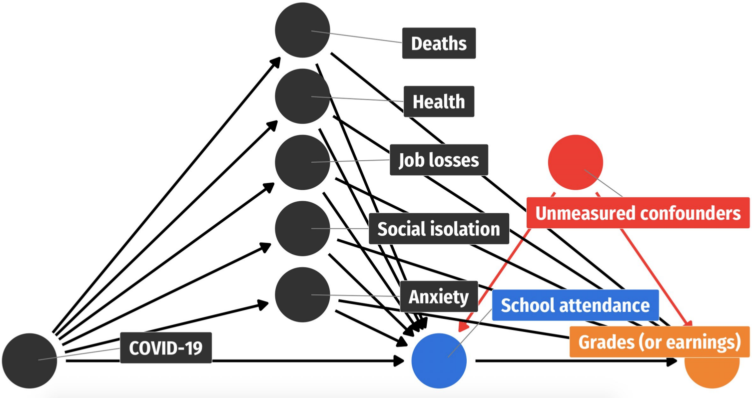
getting ready to record my instrumental variables lectures; adding obligatory "don't even think about using covid as an instrument" slides



4:18 PM · Oct 26, 2020 from Georgia, USA · Twitter for iPhone

110 Retweets **14** Quote Tweets **562** Likes

lolnope


Deaths
Health
Job losses
Unmeasured confounders
Social isolation
Anxiety
School attendance
Grades (or earnings)
COVID-19

One must justify the exclusion restriction

- You need to give good theoretical reasons why your instrument only affects the outcome through the endogenous variable
- Instrumental variables strategies live by and die from the exclusion restriction
  - Anyone who reads an IV paper will immediately try to think of ways that the exclusion restriction is violated
  - IV has become less popular as a result, because findings good instruments is tough

AJPS AMERICAN JOURNAL of POLITICAL SCIENCE

ARTICLE

Rain, rain, go away: 194 potential exclusion-restriction violations for studies using weather as an instrumental variable

Jonathan Mellon

Associate Professor, West Point Department of Systems Engineering, West Point, New York, USA

Correspondence
Jonathan Mellon, Associate Professor, West Point Department of Systems Engineering, 606 Thayer Rd, West Point, NY 10996, USA.
Email: Jonathan.mellon@westpoint.edu

Abstract
Instrumental variable (IV) analysis relies on the exclusion restriction—that the instrument only affects the dependent variable via its relationship with the independent variable and not via other causal routes. However, scholars generally justify the exclusion restriction based on its plausibility. I propose a method for searching for additional violations implied by existing social science studies. I show that the use of weather to instrument different independent variables represents strong prima facie evidence of exclusion-restriction violations for all weather-IV studies. A review of 289 studies reveals 194 variables previously linked to weather: all representing potential exclusion-restriction violations. Using sensitivity analysis, I show that the magnitude of many of these violations is sufficient to overturn numerous existing IV results. I conclude with practical steps to systematically review existing literature to identify and quantify possible exclusion-restriction violations when using IV designs.

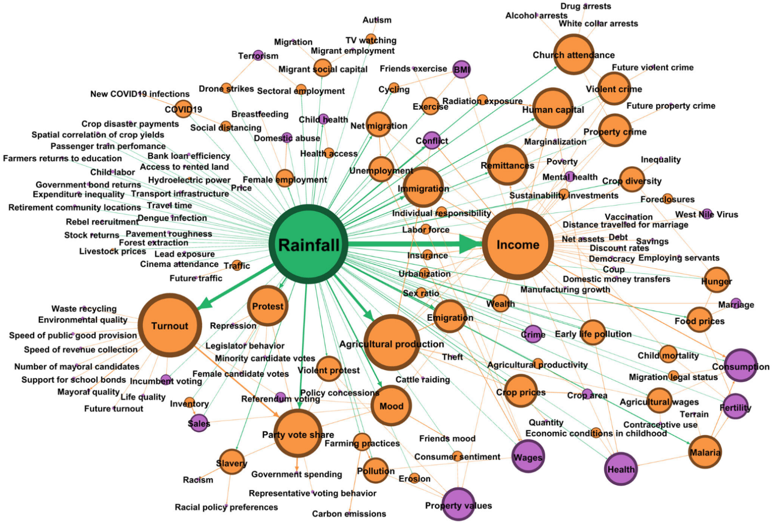


FIGURE 2 Causal web of rainfall from 192 papers. *Note*: This figure illustrates the plethora of variables causally linked to weather. These can represent exclusion-restriction violations. Node and tie sizes are proportional to appearances in literature. Colors: weather (green), instrumented variable (orange), and outcome (purple).

How do I know if I have a good instrument?

If it's weird.

But, let's say you think you do have a good instrument. How might you defend it as such to someone else? A necessary but not a sufficient condition for having an instrument that can satisfy the exclusion restriction is if people are confused when you tell them about the instrument's relationship to the outcome. Let me explain. No one is going to be confused when you tell them that you think family size will reduce female labor supply. They don't need a Becker model to convince them that women who have more children probably work less than those with fewer children. It's common sense. But, what would they think if you told them that mothers whose first two children were the same gender worked less than those whose children had a balanced sex ratio? They would probably give you a confused look. What does the gender composition of your children have to do with whether a woman works?

If it's weird, continued:

It doesn't – it only matters, in fact, if people whose first two children are the same gender decide to have a third child. Which brings us back to the original point – people buy that family size can cause women to work less, but they're confused when you say that women work less when their first two kids are the same gender. But if when you point out to them that the two children's gender induces people to have larger families than they would have otherwise, the person "gets it", then you might have an excellent instrument.

Instruments are, in other words, jarring. They're jarring precisely because of the exclusion restriction – these two things (gender composition and work) don't seem to go together. If they did go together, it would likely mean that the exclusion restriction was violated. But if they don't, then the person is confused, and that is at minimum a possible candidate for a good instrument. This is the common sense explanation of the "only through" assumption.

What would be the most ideal instrument?

- If the instrument were truly assigned at random and only operated through the endogenous variable
- We've already come across this: the Complier Average Causal Effect (CACE) in Gerber & Green (2003)
- Assignment to treatment is an instrument for receiving a door-to-door get-out-the-vote message
- We know that it causes variation in receiving a message, because the authors designed the experiment
- And the exclusion restriction almost surely holds (unless you get overly creative with other paths)

Does the exclusion restriction hold?

- Pretty easy to justify the exclusion restriction in an experiment like Gerber & Green's GOTV experiment
  - Coin flip → GOTV message

- But what about (IV → endogenous variable):
  - Rainfall → GDP growth
  - Temperature → protest
  - 2 children of the same gender → Extra child
  - Lightning → 3G internet roll-out
  - Settler mortality → Institutions
  - Soil quality → radio listenership
  - WWI war casualties → socialist support

Acemoglu et al. (2003) justify their exclusion restriction

The exclusion restriction implied by our instrumental variable regression is that, conditional on the controls included in the regression, the mortality rates of European settlers more than 100 years ago have no effect on GDP per capita today, other than their effect through institutional development. The major concern with this exclusion restriction is that the mortality rates of settlers could be correlated with the current disease environment, which may have a direct effect on economic performance. In this case, our instrumental-variables estimates may be assigning the effect of diseases on income to institutions. We believe that this is unlikely to be the case and that our exclusion restriction is plausible. The great majority of European deaths in the colonies were caused by malaria and yellow fever. Although these diseases were fatal to Europeans who had no immunity, they had limited effect on indigenous adults who had developed various types of immunities. These diseases are therefore unlikely to be the reason why many countries in Africa and Asia are very poor today (see the discussion in Section III, subsection A). This notion is

Working through an applied example: The judge IV

Misdemeanor Disenfranchisement? The Demobilizing Effects of Brief Jail Spells on Potential Voters

ARIEL WHITE *MIT*

This paper presents new causal estimates of incarceration's effect on voting, using administrative data on criminal sentencing and voter turnout. I use the random case assignment process of a major county court system as a source of exogenous variation in the sentencing of misdemeanor cases. Focusing on misdemeanor defendants allows for generalization to a large population, as such cases are very common. Among first-time misdemeanor defendants, I find evidence that receiving a short jail sentence decreases voting in the next election by several percentage points. Results differ starkly by race. White defendants show no demobilization, while Black defendants show substantial turnout decreases due to jail time. Evidence from pre-arrest voter histories suggest that this difference could be due to racial differences in exposure to arrest. These results paint a picture of large-scale, racially-disparate voter demobilization in the wake of incarceration.

Does being sent to jail for a misdemeanor cause a decrease in voter turnout?

- Doing time in jail is a memorable negative contact with the government that discourages further contact with the state
- Disrupts economic and family life
- But the type of people sent to jail are different from those who aren't
- So how can we get random variation in whether someone is sent to prison?

The naive regression results

TABLE 2. OLS Estimates of Jail's Effect on Voting

	Dependent variable		
	Voted 2012		
	(1)	(2)	(3)
Jail	−0.105*	−0.097*	−0.080*
	(0.002)	(0.002)	(0.002)
Voter birth year		−0.005*	−0.005*
		(0.0001)	(0.0001)
Black		0.115*	0.146*
		(0.002)	(0.003)
Male		−0.043*	−0.043*
		(0.002)	(0.002)
Jail × Black			−0.060*
			(0.004)
Constant	0.183*	9.466*	9.404*
	(0.001)	(0.175)	(0.174)
Observations	113,367	113,237	113,237
R^2	0.025	0.072	0.074
Adjusted R^2	0.025	0.072	0.074

Note: $^*p < 0.05$.

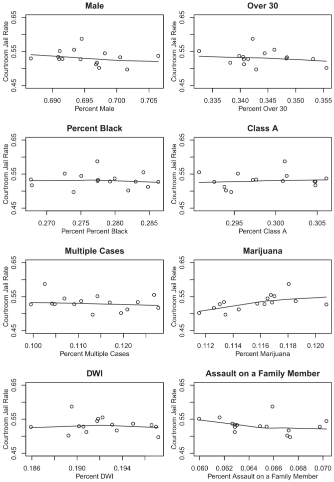
"These estimates may be biased: defendants who go to jail are probably different from those who do not go in a number of unobserved ways. But they provide a descriptive understanding of the data, and a baseline for comparison with the IV estimates."

Instrumental variables strategy

- In many US courthouses, defendants are sent at random to see one of many judges who decide cases
- Some judges are harsher in sentencing than others
- Thus being as-if randomly sent to a harsher judge increases one's probability of being sent to jail
- Judge severity → jail → voting

Is assignment to judges random?

FIGURE 1. Scatterplots of Pre-treatment Case Characteristics Against Courtroom Incarceration Rates.



Note: Each point represents one misdemeanor courtroom; lines are loess smoothers. Marijuana possession (0-2 ounces), driving while intoxicated (DWI), and assault on a family member are the most common charges in the dataset.

Model (1): First stage
Model (2): Two-stage least squares estimate

TABLE 3. Jail Sentences on 2012 Voting

	Dependent variable	
	Jail (1)	Voted 2012 (2)
Court jail average (Yr)	1.000* (0.051)	
Jail		−0.045 (0.034)
Constant	−0.0001 (0.029)	0.142* (0.019)
Year dummies	Yes	Yes
Observations	113,367	113,367
Adjusted R^2	0.004	0.017
F statistic	97.948* (df = 5; 113,361)	

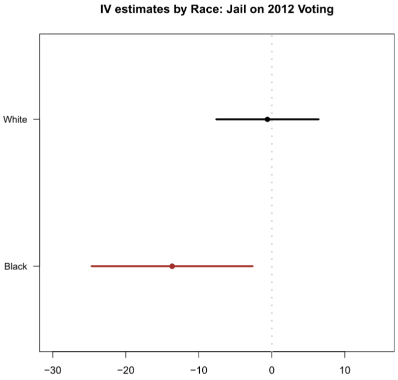
Note: $^*p < 0.05$.

Note 1: First-stage F-statistic (97.948) is well above 10

Note 2: Model (2) shows no statistically significant effect of being jailed on voting, at least when examining all defendants

Two-stage least square (TSLS) by defendant's race

FIGURE 2. Jail's Effect on Voter Turnout (2SLS Estimates), by Race of Defendant.



Note: A coefficient of −0.13 indicates a turnout decrease of 13 percentage points (among compliers).

Mechanisms? (i.e. Why does jail reduce voting)

1. Bad experience with government and lower sense of political efficacy?
   - Maybe. Cannot be tested with the data.
2. Economic and personal disruption?
   - Test differential effect on homeowners and non-homeowners. Why? Homeowners should be more shielded from resource shocks from jail
   - But author finds point estimates larger for homeowners
3. Those jailed believe that they are ineligible to vote?
   - But other research shows no difference in misinformation between those arrested and those jailed
4. In jail at time of the election? Or rearrests?
   - But very short sentences, and rearrests no more likely for non-jailed and jailed

Complete the exercise in the R file from the course website

Exercise solutions

```
# Reproduce TABLE 2
## i.e. the naive OLS assuming no confounding with controls
## Does a jail sentence mean less likely to be registered/vote?
model_1_table_2 <- lm(vote2012 ~ jail, data = D)
model_2_table_2 <- lm(vote2012 ~ jail + voterYOB + black + male, data = D)
model_3_table_2 <- lm(vote2012 ~ jail + voterYOB + black + male + jail * black,
                      data = D)
```

Exercise solutions

```
# TABLE 3 in the article
# First stage (basic OLS: linear probablity model)
# What is the first stage of the regression?
# Recall that the idea is that defendents end up, as-if random, in front of
# some judges who are harsh, and others who are lenient.
# Thus the probability that a defendent is sentenced to jail time is a
# consequence of how harsh the judge is who he or she ends up being sentenced by
# What is the outcome variable? What is the instrument?
# Note: also include "fyear" in the regression. This just controls for the year
# in which a defendent is before the judge
# Think about this regression? What is it trying to predict?
model_first_stage_table_3 <- lm(jail ~ crtjailavg1 + fyear, data = D)

# Null model
# Fit the same model, but _without_ the instrument included (i.e. just fyear)
model_first_stage_null <- lm(jail ~ fyear, data = D)
```

Exercise solutions

```
# Instrumental variables
# Here we are fitting the full instrumental variables model.
# The idea is that a person is as-if randomly assigned to a judge
# That assignment increases or decreases the probability that a defendent is put
# in jail
# Being put in jail then affects whether someone votes or not.
# Thus we are assuming the relationship is as follows:
# judge leniency (the IV) -> jail (the endogeneous variable) -> voting in 2012
# Put the first stage variables to the right of the |
# Put the second stage variables to the left of the |
# fyear should be in both the first and second stage
model_iv_table_3 <- ivreg(vote2012 ~ jail + fyear | crtjailavg1 + fyear,
                          data = D)
```

Exercise solutions

```
# TABLE A23 in the Supplementary Information & Figure 2 in the article
# The table is in the SI, but it is the main finding (the just graphs it)

# Instrumental variables regression with an interaction

# Run a first stage regression, and a 2-stage regression, as above
# But in the first stage include an interaction between
# the instrument, black, and fyear
model_first_stage_figure_2 <- lm(jail ~ crtjailavg1 * fyear * black, data = D)

# In the instrumental variables model, include an interaction between the
# instrument and "black"
model_iv_figure_2 <- ivreg(vote2012 ~ jail + fyear + black + jail * black |
                            crtjailavg1 + fyear + black + crtjailavg1 * black,
                            data = D)
```